stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 63/78 NOVEMBER
(DEPARTMENT OF MATHEMATICAL STATISTICS)

J. OOSTERHOFF

LARGE DEVIATIONS OF MULTIVARIATE L-ESTIMATORS WITH MONOTONE WEIGHT FUNCTIONS

Preprint

2e boerhaavestraat 49 amsterdam

AMS(MOS) subject classification scheme (1970): 60F10

Large deviations of multivariate L-estimators with monotone weight functions*)

by

J. Oosterhoff

ABSTRACT

A large deviation theorem for p-variate ($p \geq 1$) linear combinations of order statistics with monotone weight functions is derived. The proof is based on a Sanov-type large deviation result for empirical cdf's.

KEY WORDS & PHRASES: *large deviations, L-estimators, linear combination of order statistics.*

*) This report will be submitted for publication elsewhere.

1. INTRODUCTION AND RESULTS

Let $X_i = (X_i^{(1)},\dots,X_i^{(p)})$, $i = 1,2,\dots$ be independent p-dimensional random variables with a common distribution function (cdf) F on $\mathbb{R}^p$. Let P_G denote the probability measure on the Borel sets of $\mathbb{R}^p$ induced by a cdf G and let $\mathcal{D}^p(\mathcal{D})$ be the set of cdf's G on $\mathbb{R}^p(\mathbb{R})$ such that P_G has compact support. The Kullback-Leibler information of a cdf G w.r.t. F is defined by $K(G,F) = \int \log(dP_G/dP_F)dG$ if $P_G \ll P_F$ and $K(G,F) = \infty$ otherwise. For $\Omega \subset \mathcal{D}^p$ we write $K(\Omega,F) = \inf\{K(G,F)\colon G \in \Omega\}$; if Ω is empty $K(\Omega,F) = \infty$.

The classical result on large deviations of the multivariate sample mean is

$$(1.1) \qquad \lim_{n\to\infty} n^{-1} \log \Pr\{n^{-1} \sum_{i=1}^{n} X_i \in A\} = -K(\Omega(A),F)$$

for open convex sets $A \subset \mathbb{R}^p$, where $\Omega(A) = \{G \in \mathcal{D}^p\colon \int_{\mathbb{R}^p} x dG(x) \in A\}$. In the one-dimensional case this result is essentially due to CHERNOFF (1952). In the multivariate case (1.1) is an immediate consequence of still more general theorems in BAHADUR & ZABELL (1979) and in GROENEBOOM, OOSTERHOFF & RUYMGAART (GOR) (1979).

It is the purpose of this paper to extend this result to more general linear combinations of order statistics. Let L_1 be the space of real-valued functions on (0,1) which are integrable w.r.t. Lebesgue measure. Let $J_1,\dots,J_p \in L_1$ be weight functions and consider the statistics

$$T_n = (T_n^{(1)},\dots,T_n^{(p)}) = (\sum_{i=1}^{n} c_{i,n}^{(1)} X_{i:n}^{(1)},\dots,\sum_{i=1}^{n} c_{i,n}^{(p)} X_{i:n}^{(p)}),$$

$n \in \mathbb{N}$, where

$$c_{i,n}^{(d)} = \int_{(i-1)/n}^{i/n} J_d(u)du, \quad i = 1,\dots,n;\ d = 1,\dots,p$$

and $X_{1:n}^{(d)},\dots,X_{n:n}^{(d)}$ are the order statistics of $X_1^{(d)},\dots,X_n^{(d)}$ $(d = 1,\dots,p)$. The statistics T_n are called (multivariate) L-estimators. To avoid trivialities it will always be assumed that $J_1,\dots,J_p$ do not vanish a.e. and that P_F is nondegenerate.

It is convenient to have another representation for T_n. Let $G^{(1)},\dots,G^{(p)} \in \mathcal{D}$ denote the marginal cdf's of $G \in \mathcal{D}^p$ and let $\hat{F}_n$ be the empirical cdf of $X_1,\dots,X_n$ ($n \in \mathbb{N}$). Define the map $T_J\colon \mathcal{D}^p \longrightarrow \mathbb{R}^p$ by

$$(1.2) \qquad T_J(G) = \left(\int_0^1 J_1(u)G^{(1)^{-1}}(u)du,\dots,\int_0^1 J_p(u)G^{(p)^{-1}}(u)du\right),$$

where $G^{(d)^{-1}}(u) = \inf\{x \in \mathbb{R}\colon G^{(d)}(x) \ge u\}$, $0 < u < 1$, $d = 1,\dots,p$. Obviously

$$T_n = T_J(\hat{F}_n), \qquad n \in \mathbb{N}.$$

Let $A \subset \mathbb{R}^p$ be Borel measurable and put

$$(1.3) \qquad \Omega_J(A) = T_J^{-1}(A) = \{G \in \mathcal{D}^p\colon T_J(G) \in A\}.$$

A natural extension of (1.1) to multivariate L-estimators would be

$$(1.4) \qquad \lim_{n\to\infty} n^{-1} \log \Pr\{T_n \in A\} = -K(\Omega_J(A),F).$$

For $p = 1$ it was shown in GOR (1979) that (1.4) holds true for sets $A = [r,\infty)$ if (i) J_1 vanishes outside an interval $(\alpha,1-\alpha)$ for some $\alpha > 0$ and (ii) the map $t \to K(\Omega_J([t,\infty)),F)$ is right continuous at $t = r$.

Condition (i) above is rather restrictive. However, it turns out that this condition is redundant if J_1 is nondecreasing. Moreover, in this case condition (ii) can also be verified. The argument to prove this one-dimensional result carries over to the multivariate case.

For $A \subset \mathbb{R}^p$ let $\mathring{A}$ be its interior and $\bar{A}$ its closure in the euclidean topology. The set A is called increasing if $x \in A$ implies $y \in A$ for all $y \ge x$. Convergence of a sequence of sets $\{A_n\}$ to A in the Hausdorff metric is denoted by $A_n \to_H A$.

<u>THEOREM 1</u>. *Let $J_1,\dots,J_p \in L_1$ be nondecreasing weight functions and let $A_1,A_2,\dots \subset \mathbb{R}^p$ be Borel measurable. Then*

$$(1.5) \qquad \lim_{n\to\infty} n^{-1} \log \Pr\{T_n \in A_n\} = -K(\Omega_J(A),F)$$

if $A_n \to_H A$ *and* A *satisfies one of the following conditions:*

(C1) A *is convex and increasing and* $K(\Omega_J(\mathring{A}),F) < \infty$

(C2) A *is open,* $K(\Omega_J(A),F) = \infty$ *and* $A_n \subset A$ *for all* n.

The proof of this theorem is relegated to Section 2; it heavily leans on the methods developed in GOR (1979). The strength of the theorem lies in the fact that no smoothness conditions are imposed on the weight functions (or F). However, the condition that the weight functions be nondecreasing is quite severe. This requirement mainly serves to ensure that $T_J^{-1}(A)$ is a convex set if A is convex and increasing (cf. Lemmas 2 and 3). Obviously the theorem continues to hold if the weight functions are nonincreasing and A is convex and decreasing.

Theorem 1 implies that (1.4) holds for all open, convex and increasing sets A. This is not always true for sets A which are closed. For two-dimensional sample means a counterexample is given in GOR (1979). Although the one-dimensional Chernoff theorem holds for all intervals $[r,\infty)$, the following example demonstrates that this does not remain true for L-estimators.

EXAMPLE. Let $p = 1$ and let s_F^+ (s_F^-) denote the supremum (infimum) of the support of P_F. Let F be such that $-\infty < s_F^- < s_F^+ < \infty$, $P_F(\{s_F^-\}) > 0$ and $P_F(\{s_F^+\}) > 0$. Moreover, suppose that $J \in L_1$ is nondecreasing, changes sign on (0,1) and has at most one zero implying $\sup\{u: J(u) < 0\} = \inf\{u: J(u) > 0\} = u_0$, say $(0 < u_0 < 1)$. Define $t_0 = \sup\{T_J(G): G \in \mathcal{D}, P_G << P_F\}$ and note that $t_0 < \infty$. Then (1.4) does not hold for $A = [t_0,\infty)$. To see this, observe that $T_n \geq t_0$ iff $P_{\hat{F}_n}(\{s_F^-\}) = u_0$ and $P_{\hat{F}_n}(\{s_F^+\}) = 1 - u_0$. These equalities can only be satisfied for sample sizes n such that $nu_0 \in \mathbb{N}$. Hence $\liminf_{n\to\infty} n^{-1} \log \Pr\{T_n \geq t_0\} = -\infty$ although $K(\Omega_J([t_0,\infty)),F) < \infty$.

REMARK. In some applications the weight functions $J_1,\ldots,J_p$ also depend on n. Denoting these weight functions by $J_{1,n},\ldots,J_{p,n}$, it can be shown along the lines of the proof of Theorem 6.2 in GOR (1979) that Theorem 1 continues to hold in this case if $J_{d,n} \to J_d$ in L_1 and J_d is nondecreasing ($d = 1,\ldots,p$).

We mention the following examples of L-estimators with nondecreasing weight functions.

(a) one - or multidimensional sample means trimmed from below

(b) Gini's (one-dimensional) mean difference

$$\{n(n-1)\}^{-1} \sum_{i,j=1}^{n} |X_i - X_j| = 4n(n-1)^{-1} \int_0^1 (u-\tfrac{1}{2})\hat{F}_n^{-1}(u)du$$

(c) for $p = 1$ and cdf's F with twice differentiable densities f a well known type of L-estimator is defined by $J(u) = \psi'(F^{-1}(u))$, $0 < u < 1$, where $\psi(x) = -f'(x)/f(x)$, $x \in \mathbb{R}$ (cf. HUBER (1972)). If ψ is a convex function, J is nondecreasing and Theorem 1 is applicable. HÁJEK & ŠIDÁK (1967), p.16, mention $f(x) = \exp(x - e^{-x})$, $x \in \mathbb{R}$. Note that convexity of ψ implies that F has a thinner right-hand tail than the normal distribution.

Linear combinations of order statistics from (one-dimensional) uniform and exponential distributions can be rewritten in terms of linear combinations of i.i.d. random variables; employing this device STEINEBACH (1977) considered large deviations of such statistics.

2. PROOFS

The proof of Theorem 1 is based on a Sanov-type large deviation result from GOR (1979). Let w be the topology of weak convergence on $\mathcal{D}^p$ and let $\text{int}_w(\text{cl}_w)$ denote the interior (closure) w.r.t. this topology.

<u>LEMMA 1</u>. *Let* $\Omega \subset \mathcal{D}^p$. *If* $\{\hat{F}_n \in \Omega\}$ *is measurable* ($n \in \mathbb{N}$) *and*

$$K(\text{int}_w\, \Omega, F) = K(\text{cl}_w\, \Omega, F), \tag{2.1}$$

then

$$\lim_{n\to\infty} n^{-1} \log \Pr\{\hat{F}_n \in \Omega\} = -K(\Omega, F) \tag{2.2}$$

(see Theorem 3.1 and Lemma 2.1 in GOR (1979); a similar result is Theorem 4.5 in DONSKER & VARADHAN (1976)).

In the present notation $\{\hat{F}_n \in \Omega\}$ may be written as $\{T_n \in A\}$ and $\Omega = \Omega_J(A)$. To prove (2.2) by way of (2.1) for such sets Ω under condition (C1) of Theorem 1 is relatively easy if P_F has compact support. Superadditivity of the map $n \to \log \Pr\{T_n \in A\}$ is used to extend this result to arbitrary cdf's F. To deal with sets A_n depending on n some further continuity properties are required of $K(\Omega_J(A),F)$ as a function of A. It then remains to prove (1.5) under condition (C2), but this proof is straightforward.

We begin with some lemmas valid for nondecreasing weight functions $J, J_1, \ldots, J_p\colon (0,1) \to \mathbb{R}$. The indicator function of a set A is denoted by 1_A. Further notation will be introduced as we proceed.

LEMMA 2. *Let* $p = 1$. *If* $J \in L_1$ *is nondecreasing, the map* $T_J\colon \mathcal{D} \to \mathbb{R}$ *is concave, i.e.*

$$T_J(\alpha G_1 + (1-\alpha)G_0) \geq \alpha T_J(G_1) + (1-\alpha)T_J(G_0) \tag{2.3}$$

for all $G_0, G_1 \in \mathcal{D}$ *and all* $\alpha \in [0,1]$.

PROOF. The nondecreasing function J can be approximated by nondecreasing step functions. Since $T_J(G)$ is a linear function of J (for each fixed $G \in \mathcal{D}$), it is sufficient to prove (2.3) for functions J of the form $J = 1_{(u_0,1)}$ where $0 < u_0 < 1$ ((2.3) turns into a trivial equality if J is a constant function since $\int_0^1 G^{-1}(u)du = \int_{\mathbb{R}} x\, dG(x)$ for all $G \in \mathcal{D}$).

Let $0 < u_0 < 1$, $J = 1_{(u_0,1)}$ and $0 < \alpha < 1$. To simplify the notation we write $G_\alpha = \alpha G_1 + (1-\alpha)G_0$. We have

$$\begin{aligned}
&T_J(G_\alpha) - \alpha T_J(G_1) - (1-\alpha)T_J(G_0) = \\
&= \alpha\Big[\int_{(G_\alpha^{-1}(u_0),\infty)} y\, dG_1(y) - \int_{(G_1^{-1}(u_0),\infty)} y\, dG_1(y) + \\
&\quad + G_\alpha^{-1}(u_0)\{G_1(G_\alpha^{-1}(u_0)) - u_0\} - G_1^{-1}(u_0)\{G_1(G_1^{-1}(u_0)) - u_0\}\Big] + \\
&\quad + (1-\alpha)\Big[\int_{(G_\alpha^{-1}(u_0),\infty)} y\, dG_0(y) - \int_{(G_0^{-1}(u_0),\infty)} y\, dG_0(y) + \\
&\quad + G_\alpha^{-1}(u_0)\{G_0(G_\alpha^{-1}(u_0)) - u_0\} - G_0^{-1}(u_0)\{G_0(G_0^{-1}(u_0)) - u_0\}\Big] \\
&= \alpha V_1 + (1-\alpha)V_2, \text{ say.}
\end{aligned}$$

To show that both $V_1 \geq 0$ and $V_2 \geq 0$ first consider V_1. Three cases can be distinguished.

(a) $G_\alpha^{-1}(u_0) = G_1^{-1}(u_0)$. Obviously $V_1 = 0$ in this case.

(b) $G_\alpha^{-1}(u_0) < G_1^{-1}(u_0)$. In this case

$$V_1 = \int_{(G_\alpha^{-1}(u_0),G_1^{-1}(u_0))} y\, dG_1(y) +$$

$$+ \{G_\alpha^{-1}(u_0) - G_1^{-1}(u_0)\}\{G_1(G_1^{-1}(u_0)-) - u_0\} +$$

$$- G_\alpha^{-1}(u_0)\{G_1(G_1^{-1}(u_0)-) - G_1(G_\alpha^{-1}(u_0))\} \geq 0$$

since in the second member the first term is not smaller than the last term and the second term is nonnegative because both factors are non-positive.

(c) $G_\alpha^{-1}(u_0) > G_1^{-1}(u_0)$. By similar arguments

$$V_1 = - \int_{(G_1^{-1}(u_0),G_\alpha^{-1}(u_0)]} y\, dG_1(y) +$$

$$+ \{G_\alpha^{-1}(u_0) - G_1^{-1}(u_0)\}\{G_1(G_\alpha^{-1}(u_0)) - u_0\} +$$

$$+ G_1^{-1}(u_0)\{G_1(G_\alpha^{-1}(u_0)) - G_1(G_1^{-1}(u_0))\} \geq 0.$$

Interchanging the role of G_0 and G_1 in the above proof yields by the same argument that $V_2 \geq 0$. □

<u>LEMMA 3</u>. *Let* $J_1,\ldots,J_p \in L_1$ *be nondecreasing and let* $n = jk$ *(positive integers). Then*

(2.4) $$\Pr\{T_n \in A\} \geq (\Pr\{T_k \in A\})^j$$

for all convex increasing Borel measurable sets $A \subset \mathbb{R}^p$.

<u>PROOF</u>. Let $\hat{F}_{k,i} \in D^p$ denote the empirical cdf of $X_{(i-1)k+1},\ldots,X_{ik}$, $1 \leq i \leq j$. Repeated application of Lemma 2 to the components of the empirical cdf's yields the inequality

$$T_J(\hat{F}_n) = T_J(\sum_{i=1}^{j} j^{-1} \hat{F}_{k,i}) \geq \sum_{i=1}^{j} j^{-1} T_J(\hat{F}_{k,i}).$$

Hence the implication $T_J(\hat{F}_{k,i}) \in A$ for $i = 1,\ldots,j \Rightarrow T_J(\hat{F}_n) \in A$, for all convex increasing sets A. Since $\hat{F}_{k,1},\ldots,\hat{F}_{k,j}$ are i.i.d., (2.4) follows immediately. □

Lemma 3 expresses the superadditivity of the function $f(n) = \log \Pr\{T_n \in A\}$, $n \in \mathbb{N}$, to which we referred earlier. The proof hinges on the monotonicity of the weight functions via Lemma 2. To establish the continuity of $K(T_J^{-1}(A),F)$ as a function of A, Lemma 2 is also very convenient.

Let $e \in \mathbb{R}^p$ be the vector $(1,\ldots,1)$. For $A \subset \mathbb{R}^p$ and $x \in \mathbb{R}^p$ we write $A + x = \{y \in \mathbb{R}^p\colon\ y = z + x,\ z \in A\}$. Define the function $\kappa(\cdot;A)\colon \mathbb{R} \to \overline{\mathbb{R}}$ (the extended real line) by

$$\kappa(t;A) = K(\Omega_J(A + te),F), \qquad t \in \mathbb{R}, \tag{2.5}$$

and put $t_A = \sup\{t \in \mathbb{R}\colon \kappa(t;A) < \infty\}$. Note that $-\infty < t_A \leq \infty$.

<u>LEMMA 4</u>. *If* $J_1,\ldots,J_p \in L_1$ *are nondecreasing and* $A \subset \mathbb{R}^p$ *is a convex increasing set, the function* $\kappa(\cdot;A)$ *is continuous for all* $t \neq t_A$. *If in addition* A *is open and* $t_A < \infty$, *then* $\kappa(t_A;A) = \infty$.

<u>PROOF</u>. We first show that $\kappa(\cdot;A)$ is a convex function. Let $t_0 < t_1$, let $\varepsilon > 0$, choose $G_0 \in \mathcal{D}^p$ such that $T_J(G_0) \in A + t_0 e$ and $K(G_0,F) < \kappa(t_0;A) + \varepsilon$ and choose $G_1 \in \mathcal{D}^p$ such that $T_J(G_1) \in A + t_1 e$ and $K(G_1,F) < \kappa(t_1;A) + \varepsilon$ (if $\kappa(t_1;A) = \infty$ the proof is trivial). By Lemma 2 and the convexity of the map $s \to s \log s$, $s > 0$,

$$\begin{aligned}\kappa(\alpha t_1 + (1-\alpha)t_0;A) &\leq K(\alpha G_1 + (1-\alpha)G_0,F)\\ &\leq \alpha K(G_1,F) + (1-\alpha)K(G_0,F)\\ &< \alpha\kappa(t_1;A) + (1-\alpha)\kappa(t_0;A) + \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the convexity follows. Hence $\kappa(\cdot;A)$ is continuous at each $t \in \mathbb{R}$, $t \neq t_A$.

Now suppose A is open and $\kappa(t_A;A) < \infty$. Then there exists $G \in \mathcal{D}^p$ such that $T_J(G) \in A + t_A e$ and $K(G,F) < \infty$. Since A is open, $T_J(G) \in A + (t_A+\delta)e$ for some $\delta > 0$ implying $\kappa(t_A + \delta;A) < \infty$ in contradiction to the definition of t_A. Hence $\kappa(t_A;A) = \infty$. □

For $m = 1,2,\ldots$ let

$$B_m = \{x \in \mathbb{R}^p: \ -m \le x^{(d)} \le m,\ d = 1,\ldots,p\}.$$

Let F_m be the conditional cdf of X_1 given $X_1 \in B_m$. We write $\Pr\{T_n \in A_n | B_m\}$ to denote the conditional probability $\Pr\{T_n \in A_n \mid X_i \in B_m,\ i = 1,\ldots,n\}$. By Lemma 4.1 in GOR (1979)

$$\lim_{m\to\infty} K(\Omega,F_m) = K(\Omega,F) \quad \text{for all } \Omega \subset \mathcal{D}^p. \tag{2.6}$$

Let κ_m and $t_{A,m}$ be defined similarly to κ and t_A (cf. (2.5)), but with F replaced by F_m, $m \in \mathbb{N}$. Lemma 4 and (2.6) imply that $\lim_{m\to\infty} \kappa_m(t;A) = \kappa(t;A) < \infty$ for all $t < t_A$ if A is open, convex and increasing and hence

$$\liminf_{m\to\infty} t_{A,m} \ge t_A. \tag{2.7}$$

The map $T_{J,m}: \mathcal{D}^p \to \mathbb{R}^p$ is defined by its component functions

$$T_{J,m}^{(d)}(G) = \int_0^1 J_d(u) G^{(d)^{-1}}(u)\ 1_{B_m}(G^{(d)^{-1}}(u))du,$$

$G \in \mathcal{D}^p$, $d = 1,\ldots,p$.

<u>LEMMA 5</u>. *If* $J_1,\ldots,J_p \in L_1$, *the map* $T_{J,m}: \mathcal{D}^p \to \mathbb{R}^p$ *is continuous in the topology* w, *for all* $m \in \mathbb{N}$.

<u>PROOF</u>. Let $\{G_n\}$ be a sequence in $\mathcal{D}^p$ and suppose $G_n \to_w G$. Then $G_n^{(d)} \to_w G^{(d)}$, $d = 1,\ldots,p$, and hence $G_n^{(d)^{-1}} \to G^{(d)^{-1}}$ except in at most countably many points. Since the sequence $\{G_n^{(d)^{-1}}\ 1_{B_m}(G_n^{(d)^{-1}})\}$ is uniformly bounded, dominated convergence implies $T_{J,m}^{(d)}(G_n) \to T_{J,m}^{(d)}(G)$ for each d and thus $T_{J,m}(G_n) \to T_{J,m}(G)$. □

We are now in a position to prove Theorem 1. The proof more or less parallels the proof of Theorem 4.1 and Corollary 4.1 in GOR (1979).

PROOF OF THEOREM 1. First suppose that condition (C1) of Theorem 1 is satisfied. Since $K(\Omega_J(\mathring{A}),F) < \infty$, Lemma 4 implies that $t_{\mathring{A}} > 0$. Let $-\infty < \eta < t_{\mathring{A}}$. It follows from (2.7) that $\kappa_m(\eta;\mathring{A}) < \infty$ for all $m \geq m_\eta$. Choose $\varepsilon > 0$ and fix $m \geq m_\eta$. Let $G_0 \in \Omega_J(\bar{A}+\eta e)$ be such that $K(G_0,F_m) < \kappa_m(\eta;\bar{A}) + \varepsilon$ and let $G_1 \in \Omega_J(\mathring{A}+\eta e)$ be such that $K(G_1,F_m) < \infty$. Put $G_\alpha = \alpha G_1 + (1-\alpha)G_0$, $0<\alpha<1$. Then $G_\alpha \in \Omega_J(\mathring{A}+\eta e)$ by Lemma 2 and hence by the convexity of the map $G \to K(G,F_m)$

$$\kappa_m(\eta;\mathring{A}) \leq \lim_{\alpha\downarrow 0} K(G_\alpha,F_m) \leq$$

$$\leq \lim_{\alpha\downarrow 0} \{\alpha K(G_1,F_m) + (1-\alpha)K(G_0,F_m)\} = K(G_0,F_m) <$$

$$< \kappa_m(\eta;\bar{A}) + \varepsilon .$$

It follows that for $m \geq m_\eta$

$$K(\Omega_J(\mathring{A}+\eta e),F_m) = K(\Omega_J(\bar{A}+\eta e),F_m).$$

This equality continues to hold if the sets $\Omega_J(\mathring{A}+\eta e)$ and $\Omega_J(\bar{A}+\eta e)$ are replaced by $T_{J,m}^{-1}(\mathring{A}+\eta e)$ and $T_{J,m}^{-1}(\bar{A}+\eta e)$, respectively. By Lemma 5 $T_{J,m}^{-1}(\mathring{A}+\eta e)$ is w-open and $T_{J,m}^{-1}(\bar{A}+\eta e)$ is w-closed. Since, moreover,

$$\Pr\{T_J(\hat{F}_n) = T_{J,m}(\hat{F}_n) \mid B_m\} = 1$$

for all m and n, Lemma 1 implies

$$\lim_{n\to\infty} n^{-1} \log \Pr\{T_n \in \bar{A}+\eta e \mid B_m\} = -K(\Omega_J(A+\eta e),F_m) \tag{2.8}$$

for all $m \geq m_\eta$ and $\eta < t_{\mathring{A}}$ (note that $\operatorname{int}\bar{A} = \mathring{A}$ since A is convex).

Fix $\eta < 0$, let $\varepsilon > 0$, define

$$\pi_\eta = \limsup_{n\to\infty} n^{-1} \log \Pr\{T_n \in \bar{A}+\eta e\}$$

and choose $k = k(\varepsilon,\eta) \in \mathbb{N}$ so large that

$$\pi_\eta \leq k^{-1} \log \Pr\{T_k \in \bar{A}+\eta e\} + \varepsilon .$$

Since $\lim_{m\to\infty} \Pr\{T_k \in \bar{A} + \eta e | B_m\} = \Pr\{T_k \in \bar{A} + \eta e\}$, there exists $m_{\eta,k} \geq m_\eta$ such that

$$\pi_\eta \leq k^{-1} \log \Pr\{T_k \in \bar{A} + \eta e | B_m\} + 2\varepsilon$$

for all $m \geq m_{\eta,k}$. Furthermore, by Lemma 3

$$\limsup_{n\to\infty} n^{-1} \log \Pr\{T_n \in \bar{A} + \eta e | B_m\}$$

$$\geq \lim_{j\to\infty} (kj)^{-1} \log(\Pr\{T_k \in \bar{A} + \eta e | B_m\})^j$$

$$= k^{-1} \log \Pr\{T_k \in \bar{A} + \eta e | B_m\}$$

for all $m \in \mathbb{N}$, and hence in view of (2.8)

$$\pi_\eta \leq -K(\Omega_J(A + \eta e), F_m) + 2\varepsilon$$

for $m \geq m_{\eta,k}$. Since $A_n \to_H A$ implies $A_n \subset \bar{A} + \eta e$ ($\eta < 0$) for sufficiently large n, it follows from the preceding inequality and (2.6) that

$$\limsup_{n\to\infty} n^{-1} \log \Pr\{T_n \in A_n\} \leq \pi_\eta \leq -K(\Omega_J(A + \eta e), F)$$

for all $\eta < 0$. Application of Lemma 4 yields

$$\limsup_{n\to\infty} n^{-1} \log \Pr\{T_n \in A_n\} \leq -K(\Omega_J(A), F). \tag{2.9}$$

Conversely, for any $m,n \in \mathbb{N}$

$$n^{-1} \log \Pr\{T_n \in A_n\} \geq n^{-1} \log \Pr\{T_n \in A_n | B_m\} + \log P_F(B_m).$$

Choosing $0 < \eta < t_{\mathring{A}}$, it follows by (2.8) and (2.6)

$$\liminf_{n\to\infty} n^{-1} \log \Pr\{T_n \in A_n\}$$

$$\geq \liminf_{m\to\infty}[\liminf_{n\to\infty} n^{-1} \log \Pr\{T_n \in A_n | B_m\} + \log P_F(B_m)] \geq$$

$$\geq \liminf_{m\to\infty}[\liminf_{n\to\infty} n^{-1} \log \Pr\{T_n \in \bar{A} + \eta e | B_m\} + \log P_F(B_m)]$$

$$= - \lim_{m\to\infty} K(\Omega_J(A+\eta e), F_m) = - K(\Omega_J(A+\eta e), F).$$

Lemma 4 again implies

$$\liminf_{n\to\infty} n^{-1} \log \Pr\{T_n \in A_n\} \geq - K(\Omega_J(A), F).$$

In combination with (2.9) this establishes (1.5) under condition (C1).

Now let $A \subset \mathbb{R}^p$ be open and $K(\Omega_J(A),F) = \infty$. Put $\Psi_m = \{G \in \mathcal{D}^p: P_G$ has support in $B_m\}$, $m \in \mathbb{N}$. Suppose there exists $G \in \Psi_m$ such that P_G has finite support contained in the support of P_{F_m} and such that $T_J(G) = T_{J,m}(G) \in A$. Since $T_{J,m}$ is a w-continuous map (Lemma 5), there exists a w-open neighborhood V_G of G such that $T_{J,m}(V_G) \subset A$. Lemma 2.5 in GOR (1979) states that $G \in \mathrm{cl}_w\{H \in \mathcal{D}^p: K(H,F_m) < \infty\}$ and hence there exists $G_V \in V_G$ such that $K(G_V,F_m) < \infty$. It follows that $G_V \in \Psi_m \cap V_G$, implying $T_J(G_V) = T_{J,m}(G_V) \in A$ and hence $K(\Omega_J(A),F_m) < \infty$ in contradiction to the original assumption. But this means that $T_J(G) \notin A$ if P_G has finite support in the support of P_{F_m}, i.e.

$$\Pr\{T_J(\hat{F}_n) \in A | B_m\} = 0$$

for all $m,n \in \mathbb{N}$. Hence $\Pr\{T_n \in A\} = 0$ and together with $A_n \subset A$ this implies $\Pr\{T_n \in A_n\} = 0$ for all n. This proves (1.5) under condition (C2). □

REFERENCES

[1] BAHADUR, R.R. & S.L. ZABELL (1979), *Large deviations of the sample mean in general vector spaces*, Ann. Prob. 7, to appear.

[2] CHERNOFF, H. (1952), *A measure of asymptotic efficiency of tests of a hypothesis based on sums of observations*, Ann. Math. Statist. 23, 493-507.

[3] DONSKER, M.D. & S.R.S. VARADHAN (1976), *Asymptotic evaluation of certain Markov process expectations for large time* III, Comm. Pure Appl. Math. 29, 389-461.

[4] GROENEBOOM, P., OOSTERHOFF, J. & F.H. RUYMGAART (1979), *Large deviation theorems for empirical probability measures*, Ann. Prob. 7, to appear.

[5] HÁJEK, J. & Z. ŠIDÁK (1967), *Theory of rank tests*, Academic Press, New York.

[6] HUBER, P. (1972), *Robust statistics: a review*, Ann. Math. Statist. 43, 1041-1067.

[7] STEINEBACH, J. (1977), *Exponential convergence properties of linear estimators under exponential and uniform distribution*, Metrika 24, 137-161.